Human-Centered AI

Abstract

We identify three specific areas of focus to advance human-centered AI:• Designers and systems must understand the context of use and sense changes over time: Successful AI Engineering depends on the team's ability to identify and articulate the desired system outcome and understand human and contextual factors affecting the outcome. The system itself must be able to learn when shifts in context have occurred. What are the best ways to maintain clarity around operational intent and mechanisms for adapting and evolving systems based on dynamic contexts and user needs?• Development of tools, processes, and practices to scope and facilitate human-machine teaming: Implementation of AI systems entails high levels of interdependence between human and machine. Adoption of AI systems requires the primary users to interact with and understand systems, gaining appropriate levels of trust. Every AI system needs to be designed to recognize boundaries and unfamiliar scenarios, and to provide transparency regarding its limitations.• Methods, mechanisms, and mindsets to engage in critical oversight: AI systems learn through data and observations, rather than being explicitly programmed for a deterministic outcome. Critical and reflective oversight by organizations, teams, and individuals that create and use AI systems is needed to uphold ethical principles and proactively consider the risks of bias, misuse, abuse, and unintended consequences through design, development, and ongoing deployment.For each area, we identify ongoing work as well and challenges and opportunities in developing and deploying AI systems with confidence.